

On the Impact of Optimal Modulation and FEC Overhead on Future Optical Networks

Alex Alvarado, David J. Ives, Seb Savory and Polina Bayvel

Abstract—The potential of optimum selection of modulation and forward error correction (FEC) overhead (OH) in future transparent nonlinear optical mesh networks is studied from an information theory perspective. Different network topologies are studied as well as both ideal soft-decision (SD) and hard-decision (HD) FEC based on demap-and-decode (bit-wise) receivers. When compared to the de-facto QPSK with 7% OH, our results show large gains in network throughput. When compared to SD-FEC, HD-FEC is shown to cause network throughput losses of 12%, 15%, and 20% for a country, continental, and global network topology, respectively. Furthermore, it is shown that most of the theoretically possible gains can be achieved by using one modulation format and only two OHs. This is in contrast to the infinite number of OHs required in the ideal case. The obtained optimal OHs are between 5% and 80%, which highlights the potential advantage of using FEC with high OHs.

Index Terms—Bit-wise receivers, channel coding, forward error correction, modulation, soft-decision, optical networks.

I. INTRODUCTION AND MOTIVATION

The rapid rise in the use of the Internet has led to increasing traffic demands putting severe pressure on backbone networks. The transport backbone of the Internet is formed of optical mesh networks, where optical fibre links connect nodes formed of reconfigurable optical add drop multiplexers (ROADMs). Studying the ultimate transmission limits of optical mesh networks as well as the optimal utilization of the installed network resources is therefore key to avoid the so-called “capacity crunch” [1], [2].

Installed optical mesh networks utilize wavelength routing to transparently connect source and destination transceivers. The quality of the optical communication signal degrades due to transmission impairments, which in turn limits the maximum achievable data rate. This degradation is usually due to the amplifiers in the link as well as nonlinear distortion due to neighboring WDM channels. Furthermore, the optical signals within a transparent wavelength routed network travel a variety of distances, and thus, experience different levels of signal degradation. The most conservative design alternative for an optical network is to choose the transceiver to operate error-free on the worst light path, i.e., for transmission between the furthest spaced nodes [3, p. 138], [4, Sec. 1]. Under this design

Research supported by Engineering and Physical Sciences Research Council (EPSRC) through the CDT in Photonic Systems Development and the project UNLOC (EP/J017582/1), and by the Royal Academy of Engineering/The Leverhulme Trust Senior Research Fellowship. This work was presented in part at the 2015 Optical Fiber Communication Conference (OFC), Los Angeles, CA, Mar. 2015.

The authors are with the Optical Networks Group, Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, United Kingdom (email: alex.alvarado@ieee.org).

paradigm, any route reconfiguration can be accommodated through the network. However, this leads to over provisioning of resources.

The increase in traffic demand together with the development of software-defined transceivers that can adapt the transmission parameters to the physical channel have increased the interest in designing networks that utilize the resources more efficiently. The degrees of freedom in the transceiver include, e.g., the forward error correction (FEC) scheme, FEC overhead (OH), modulation format, frequency separation (in flex-grid networks), launch power, and symbol rate (see [5, Fig. 3]). The network resources can be better utilized if these degrees of freedom are jointly optimized in conjunction with the routing of the optical light path through the network.

To cope with increasing capacity demands, future optical networks will use multilevel modulations and FEC. This combination is known as coded modulation (CM) and its design requires the joint optimization of the FEC and modulation format (see Fig. 1). The optimum receiver structure for CM is the maximum likelihood (ML) receiver, which finds the most likely *coded sequence* [6, Sec. 3.1]. The ML solution is in general impractical, and thus, very often the receiver is implemented as a (suboptimal) bit-wise (BW) receiver instead [6, Sec. 3.2]. In a BW receiver, hard or soft information on the code bits is calculated first, and then, an FEC decoder is used (see the receiver side of Fig. 1). In other words, practical receivers decouple the detection process: symbols are first converted into bits, and then, FEC decoding is applied. In this paper we consider BW receivers with both soft-decision FEC (SD-FEC) and hard-decision FEC (HD-FEC). BW receivers have been studied for optical communications in, e.g., [7], [8], [9].

Traditional analyses of optical networks assume a target pre-FEC bit error rate (BER), and thus, an HD-FEC with fixed OH is implicitly assumed.¹ Although the most common value for the OH is 7%, higher OHs have become increasingly popular. Furthermore, the use of SD-FEC with high OHs (typically around 20%) is considered the most promising FEC alternative for 100G and 400G transceivers. Although from a theoretical point of view, fixing the OH is an artificial constraint that reduces flexibility of the CM design and reduces the network throughput, there are good reasons for fixing the OH. The client rates are usually quantized to 10, 40, 100, 400 Gb/s for compatibility with the Ethernet standards. The symbol rate is also often fixed to accommodate the transmitted bandwidth

¹When SD-FEC is considered, however, pre-FEC BER does not determine the FEC OH, as recently shown in [10].

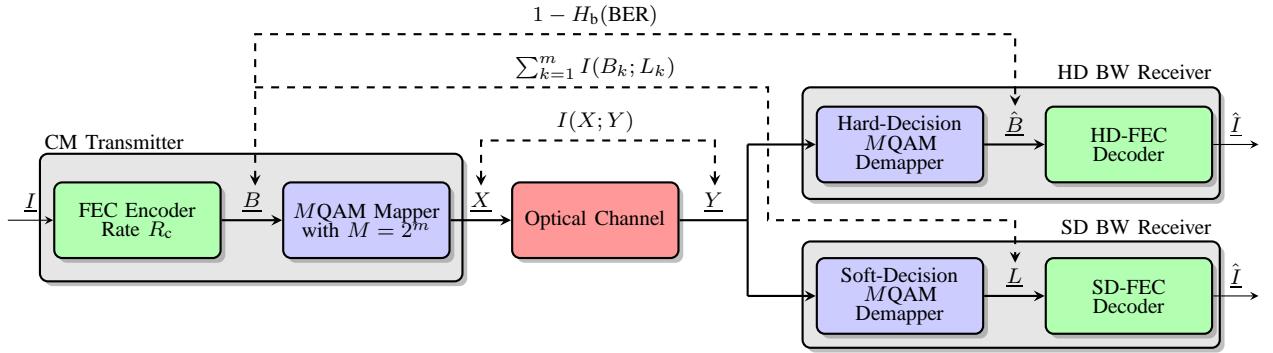


Fig. 1. Coded modulation system under consideration. At the transmitter, a binary FEC code is concatenated with a M -ary QAM modulator. After transmission, the noisy received symbols are demapped and then decoded by a bit-wise receiver. When the demapper makes hard decisions on the symbols, an HD-FEC is used. When the demapper computes log-likelihood ratios (soft-decisions), an SD-FEC is assumed. The achievable rates discussed in Sec. III are also shown.

into a given fixed-grid, leading to fixed OHs. Under these constraints, no full flexibility on the selection of OHs is possible. In this paper, however, we will ignore these constraints and focus on finding the theoretically maximum network throughput.

To increase the network throughput, different approaches have been investigated in the literature. For example, [11] considered mixed line rates (10, 40, 100 Gb/s) for the NSF mesh topology, [12] considered variable FEC OHs with fixed symbol rate and modulation formats, and [13] considered variable modulation format and SD-FEC OHs. Adaptive FEC based on practical codes was recently considered in [4]. Variable OHs with 16QAM and 64QAM were studied in [14], where probabilistically-shaped constellations were considered. The optimal modulation format based on an approximation for the maximum achievable rates of HD-FEC was considered in [15]. Adaptive FEC OH for time frequency packing transmission was studied in [16] and [17]. The problem of routing and spectrum assignment for flex-grid optical networks with orthogonal frequency-division multiplexing was studied in [18]. The key enabling technology for these approaches are software-defined transceivers, allowing for example to vary the modulation format and symbol rate, as done in [19], [20], [21], [22], [23]. In [24], [17], a software defined transceiver with variable FEC was experimentally demonstrated.

To optimize the network design, a physical layer model is required. While in the past very simple models (e.g., reach-based models) were considered, recently, nonlinear effects have been taken into account using the Gaussian noise (GN) model [25], [26], [27], [28], [29], [30], [31]. In [32], the closed form solution of the GN model of [31] was used to adapt the routing and wavelength assignment problem, for a target pre-FEC BER, and 4 different modulation formats. The same GN model was used in [33] and [34] to jointly optimize power, modulation format, and carrier frequencies (flex-grid) for a fixed OH. Numerical integration of the GN model was used in [35] to assess SNR and throughput optimization via power and modulation adaptation. The numerically integrated GN model in [35] was also used in [36], [37] to sequentially optimize modulation format and power for a fixed FEC OH. The potential gains of adaptive FEC OH and modulation, or adaptive launch power and modulation were studied in [38].

In this paper, we study the problem of finding the optimal modulation and FEC OH from an information theory viewpoint. In particular, we use information theoretic quantities (i.e., achievable rates) and a realistic model for the nonlinear interference to study the maximum network throughput of optical mesh networks. For point-to-point links, and under a Gaussian assumption on the channel, the solution depends only on the signal-to-noise ratio (SNR). For an optical network, however, the solution depends on the SNR distribution of the connections. Therefore, the theoretically optimum CM design is obtained when the modulation size and FEC OH are jointly designed across the network. Significant increases in network throughput are shown. Furthermore, practically relevant schemes (based on either one or two OHs) are also considered, and their gap to the theoretical maximum is quantified. This paper extends our results in [39] by considering multiple network topologies as well as both HD- and SD-FEC.

This paper is organized as follows. In Sec. II the system model, network topologies, and physical layer model are described. In Sec. III the optimal selection of modulation and coding is reviewed and the maximum network throughput is analyzed. In Sec. IV practically relevant schemes are considered. Conclusions are drawn in Sec. V.

II. PRELIMINARIES

A. System Model

We consider the CM transmitter shown in Fig. 1, where a binary FEC code maps the *information bits* $\underline{I} = [I_1, I_2, \dots, I_{k_c}]$ into a sequence of code bits $\underline{B} = [B_1, B_2, \dots, B_{n_c}]$, where $R_c = k_c/n_c$ is the code rate. At each discrete-time instant, m code bits are mapped to a constellation symbols from a discrete constellation with $M = 2^m$ constellation points. We consider polarization-multiplexed square QAM constellations with $m = 2, 4, 6, 8, 10$ (i.e., MQAM with $M \leq 1024$) in each polarization. The constellations are labeled by the binary-reflected Gray code. For FEC encoder with code rate R_c and a modulation format with M symbols, the spectral efficiency (SE) per two polarizations is

$$\text{SE} = 2R_c m \left[\frac{\text{bit}}{\text{symbol}} \right]. \quad (1)$$

The FEC OH is

$$\text{OH} = 100 \left(\frac{1}{R_c} - 1 \right) \%. \quad (2)$$

At the receiver side a bit-wise (BW) receiver is used². In such a receiver, the noisy symbols are first demapped, and then, FEC decoding is performed. The FEC decoder gives an estimate of the transmitted bits \hat{L} . Due to the separation of the detection process, BW receivers are suboptimal. They are, however, very popular in practice due to the use of off-the-shelf FEC decoders.

In this paper we consider two types of FEC decoders: hard- and soft-decision FEC. This naturally leads to two different BW receiver structures, shown on the right-hand side of Fig. 1. In a hard-decision BW receiver, the demapper makes hard decision on the bits (by making hard-decision on the symbols), which are then passed to an HD-FEC decoder. We assume that there is a random bit-level interleaver between the encoder and mapper, which we consider part of the FEC encoder (decoder).

In a soft-decision BW receiver, the demapper calculates *soft* information on the code bits (also known as “soft bits”), which are then passed to an SD-FEC decoder. The soft information is usually represented in the form of logarithmic likelihood ratios (LLRs), defined as

$$L_q \triangleq \log \frac{f_{Y|B_q}(y|1)}{f_{Y|B_q}(y|0)}, q = 1, \dots, m, \quad (3)$$

where B_q is the q th bit at the input of the mapper, and $f_{Y|B_q}(y|b)$, $b \in \{0, 1\}$ is the channel transition probability. The sign of the LLRs corresponds to a hard-decision on the bits, while its amplitude represents the reliability of the available information.

B. Network Topologies

In this paper we consider 3 networks shown in Figs. 2, 3, and 4. The first one is the exemplary network topology for Deutsche Telekom Germany (DTG) [40, Sec. II], where the two core nodes per city (see [40, Fig. 1]) were merged into one. The second one is the reference 14-node, 21-link NSF mesh topology [36, Fig. 1], [35, Fig. 7]. The last topology is the Google B4 (GB4) network connecting data centers in [41, Fig. 1]. We chose to study these three topologies because they are representative of networks at three different scales: country, continental, and global.

C. Physical Layer Model

For the analysis in this paper, it is assumed that each node in the networks described in Sec. II-B is equipped with multiple transceivers. Furthermore, it is also assumed that these transceivers are based on the structure in Fig. 1 and that they can ideally adapt the FEC OH and modulation order. We consider a fixed grid of 80 WDM channels of 50 GHz and a symbol rate of 32 GBaud.

The nodes are connected with fiber pairs with standard single-mode fiber with parameters shown in Table I. Erbium-doped fiber amplifiers (EDFA) are regularly placed between

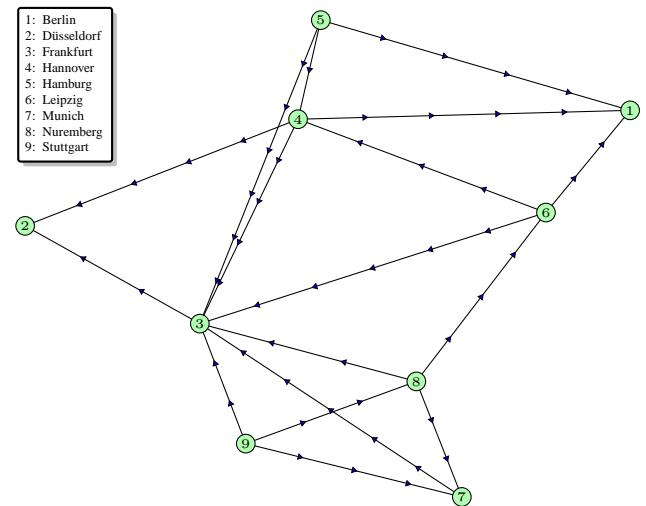


Fig. 2. DTG network formed by 9 nodes representative of a country-level topology. The location of the amplifiers are shown as triangles.

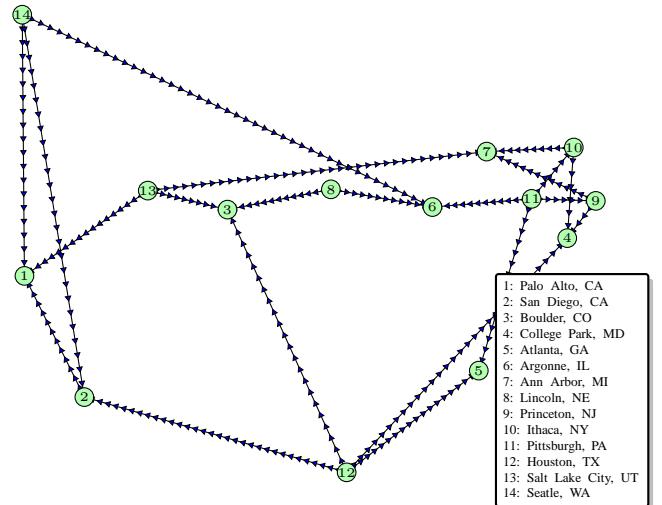


Fig. 3. NSF network formed by 14 nodes representative of a continental-level topology. The location of the amplifiers are shown as triangles.

the links, as shown in Figs. 2 and 3³. The span length is 80 km and the EDFA noise figure is 5 dB.

Every link in the network is modeled using a channel model (see Fig. 1) that encompasses all the transmitter digital signal processing (DSP) used after the mapper (i.e., pulse shaping, polarization multiplexing, filtering, electro-optical conversion, etc.), the physical channel (the fiber and amplifiers), and the receiver DSP (optical-to-electrical conversion, filtering, equalization, matched filtering, etc.). This channel is modeled using a dual-polarization, discrete-time, memoryless, additive white Gaussian noise (AWGN) channel.

For each polarization, and at each discrete time $k = 1, 2, \dots, n$, $Y_k = X_k + Z_k$, where X_k is the transmitted symbol, Z_k are independent, zero-mean, circularly symmetric, complex Gaussian random variables, and n is the blocklength. This GN channel model characterizes optically-amplified links dominated by amplified spontaneous emission noise where

²Also known as bit-interleaved coded modulation receiver [6]

³The amplifiers are not shown in Fig. 4 so as not to overcrowd the figure.

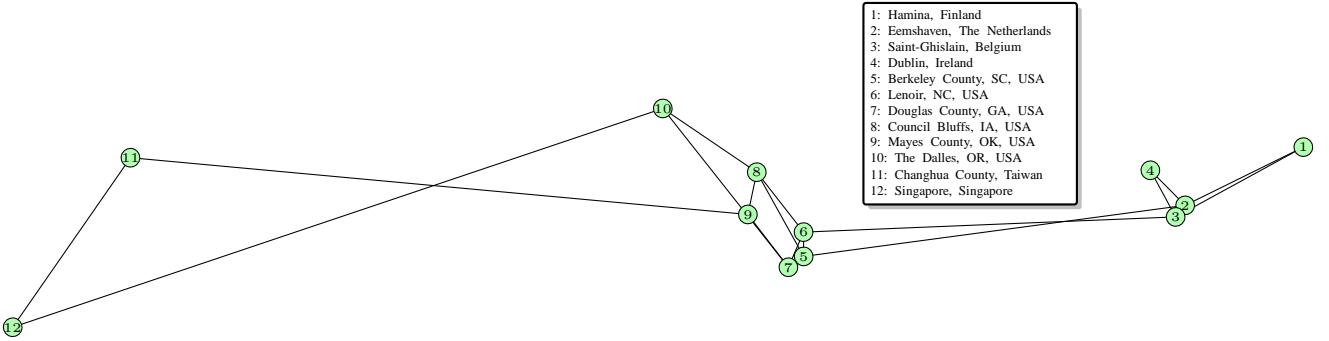


Fig. 4. GB4 network (amplifiers are not shown) formed by 12 nodes representative of a global-level topology.

TABLE I
SUMMARY OF SYSTEM PARAMETERS.

Parameter	Value
Fiber attenuation	0.22 dB/km
Dispersion parameter	16.7 ps/nm/km
Fiber nonlinear coefficient	1.3 1/W/km
Span length	80 km
EDFA noise figure	5 dB
ASE Noise per span P_{ASE}	0.747 μW
Symbol rate	32 Gbaud
WDM channels	80
Channel separation	50 GHz
Pulse shape	Nyquist sinc pulses
SPM	Ideally compensated
ROADM Loss	0 dB
Nonlinear coefficient η	742 1/W ²
Launch power per channel P	-1 dBm

chromatic dispersion and polarization mode dispersion are perfectly compensated. In this model, the power-dependent nonlinearities in the presence of sufficient dispersion are treated as an additional source of AWGN.

The symbol SNR for a route with N_s spans is calculated as

$$\text{SNR} = \frac{P}{N_s(P_{\text{ASE}} + \eta P^3)} \quad (4)$$

where P is the launch power per channel, P_{ASE} is the ASE noise added after each span (in the 32 GHz signal bandwidth), η is the nonlinear coefficient (per span).

The nonlinear interference is taken as that on the worst case central DWDM channel, i.e., we assumed that all the links were fully loaded with DWDM channels. The nonlinear coefficient η is calculated using the incoherent GN model of nonlinear interference [30], SPM is assumed to be ideally compensated, and the ROADM nodes were assumed lossless. Using the parameters in Table I, we obtain $\eta \approx 742 \text{ W}^{-2}$. The launch power that maximizes the SNR in (4) is found to be -1 dBm. A summary of the parameters discussed in this section is given in Table I.

D. Performance Metric

Throughout this paper, the main performance metric considered is the total *network throughput*, which we define below

and denote by Θ . The network throughput is the total traffic transported by the network that satisfies the required traffic profile.

The network is assumed to have N nodes. The required connectivity is defined by the normalized traffic profile $T_{s,d}$ with $s, d \in \mathcal{N}$, where $\mathcal{N} \triangleq \{1, \dots, N\}$ is the set of network nodes. The traffic profile is normalized such that

$$\sum_{s=1}^N \sum_{d=1}^N T_{s,d} = 1. \quad (5)$$

We assume connectivity between all pairs of nodes is required ($0 < T_{s,d} < 1, s \neq d$) and that $T_{s,d} = 0$ if $s = d$.

Let $C_{s,d}$ be the total available throughput between nodes s and d (across different routes), where

$$C_{s,d} \triangleq \sum_{r=1}^{R_{s,d}} C_{s,d}^{(r)} \quad (6)$$

where $C_{s,d}^{(r)}$ is the available throughput in the r th route, and $R_{s,d}$ is the number of active routes between nodes s and d . The network throughput Θ is then defined as

$$\Theta \triangleq \min_{\substack{s, d \in \mathcal{N} \\ d \neq s}} \frac{C_{s,d}}{T_{s,d}}. \quad (7)$$

In this paper, we consider a uniform traffic profile, i.e.,

$$T_{s,d} = \begin{cases} \frac{1}{N(N-1)}, & s \neq d \\ 0, & s = d \end{cases}. \quad (8)$$

The network throughput in this case can be expressed as

$$\Theta = N(N-1) \min_{\substack{s, d \in \mathcal{N} \\ d \neq s}} C_{s,d} \quad (9)$$

$$= N(N-1) \min_{\substack{s, d \in \mathcal{N} \\ d \neq s}} \sum_{r=1}^{R_{s,d}} C_{s,d}^{(r)} \quad (10)$$

where (10) follows from (6).

III. OPTIMAL MODULATION AND FEC OVERHEAD

In this section, we describe the selection of optimal FEC OH and modulation format from an information theory viewpoint. We first consider the ideal case of continuous constellations and then move to the case of modulation with discrete number

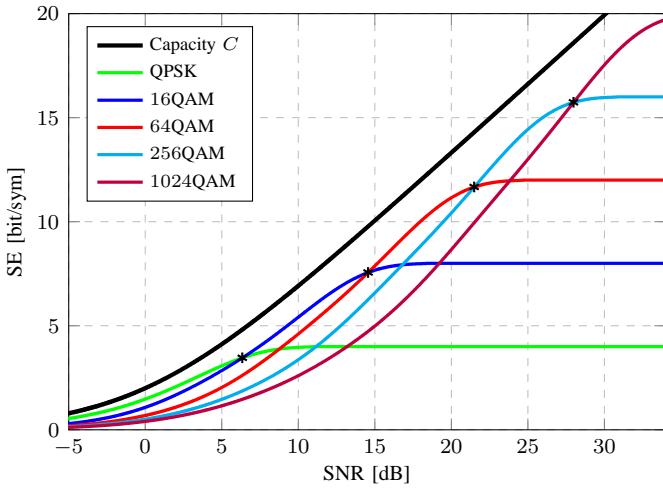


Fig. 5. Achievable rates for a BW receiver with HD-FEC. The black asterisks show the SNR values where the modulation size should be changed. The channel capacity in (11) is shown for comparison.

of constellation points. The routing and wavelength assignment problem and the maximum network throughput are also discussed in this section.

A. Channel Capacity

The capacity of the AWGN channel (in [bit/sym]) under an average power constraint is

$$C = 2 \log_2(1 + \text{SNR}) \quad (11)$$

where the pre-log factor of 2 takes into account the two polarizations. The value of C represents the maximum number of information bits per symbol that can be reliably transmitted through an AWGN channel.

The capacity in (11) is achieved when the transmitted symbols are chosen from a zero-mean Gaussian distribution. In practice, however, the modulation has M discrete levels, which reduces the achievable transmission rates. This case is studied in the next section.

B. Achievable Rates for Discrete Constellations

From an information theory point of view, the optimal code rate and constellation size can be chosen from the *mutual information* (MI). The MI, usually denoted by $I(X; Y)$, is an achievable rate for an optimal receiver.⁴ MI curves for square QAM constellations indicate that, regardless of the SNR, in order to maximize the SE, the densest available constellation should always be used and the code rate chosen between 0 and 1.⁵ This has been shown, e.g., in [6, Fig. 4.3], [14, Fig. 1], [30, Fig. 11].

The MI is not an achievable rate for the two receiver structures we consider in this paper (see Fig. 1). The first

⁴In information theory, a transmission rate R is said to be achievable if there exist an encoder with rate R and a (possibly very complex) decoder, such that the error probability after decoding vanishes as $n \rightarrow \infty$.

⁵In practice, of course, this is difficult as using very dense constellations at low SNRs are hard to realize.

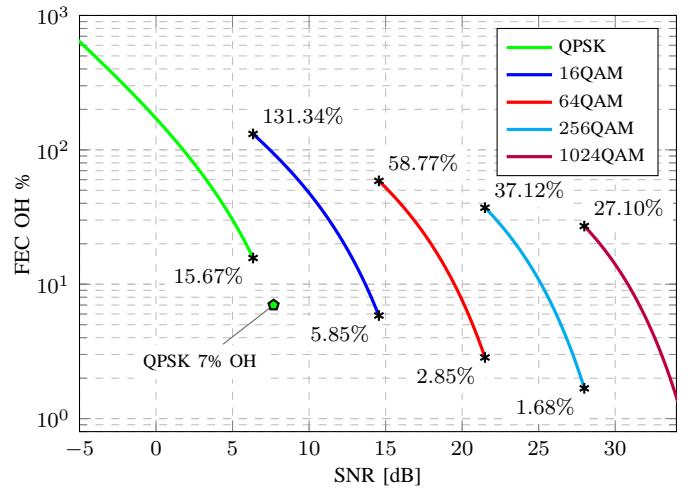


Fig. 6. OHs obtained from the achievable rates in Fig. 5 for a BW receiver with HD-FEC. The black asterisks show the SNR values where the modulation size should be changed and the corresponding FEC OH. The channel capacity in (11) is shown for comparison. The SNR required for QPSK with 7% FEC OH is also shown with a green pentagon.

receiver in Fig. 1 is suboptimal because it makes hard-decisions on the symbols (and thus, information is lost). The second receiver is suboptimal because the LLR calculation ignores the dependency of the bits within a symbol (i.e., L_q in (3) does not depend on B_l for $l \neq q$). In this paper we consider two different achievable rates, one for each of these receiver structures.

For the case of a BW receiver with hard-decisions, Shannon's coding theorem state that error-free transmission is possible when $n \rightarrow \infty$ if the rate of the encoder fulfills [15, eq. (5)] [42, eq. (8)]

$$R_c \leq \frac{I(B; \hat{B})}{m} = 1 - H_b(\text{BER}) \quad (12)$$

where $I(B; \hat{B})$ is the MI between the transmitted and received code bits (see Fig. 1). In (12), $H_b(p) = -p \log_2 p - (1-p) \log_2 (1-p)$ is the binary entropy function and BER is the average pre-FEC BER (across m bit positions).

The values of $I(B; \hat{B})$ for $M = 4, 16, \dots, 1024$ are shown in Fig. 5, where the BER was calculated using [43, Th. 2]. The key difference between the achievable rates $I(B; \hat{B})$ and $I(X; Y)$ is that the former cross each other for different values of M (see the asterisks in Fig. 5). An important consequence of crossing achievable rate curves is that the theoretically optimal choice of R_c and M is not straightforward. For a BW receiver with HD-FEC we consider here, QPSK should be used for SNRs below $\text{SNR} \leq 5.8$ dB, 16QAM for $5.8 \leq \text{SNR} \leq 14$ dB, etc. The corresponding FEC OHs obtained via (2) are shown in Fig. 8. This figure also shows the optimum minimum and maximum OH values for each modulation format as well as the SNR required for QPSK with 7% FEC OH (green pentagon).

When the BW receiver operates with soft-decisions, and if the code bits are independent, an achievable rate is given by

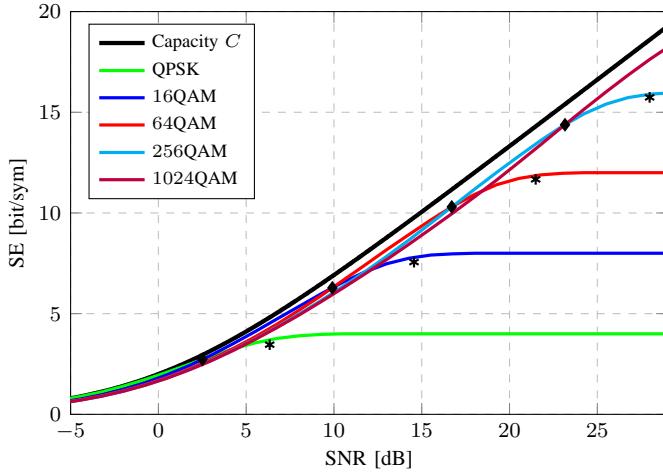


Fig. 7. Achievable rates (GMI) for a BW receiver with SD-FEC. The black diamonds show the SNR values where the modulation size should be changed. The channel capacity in (11) and the crossing points of the achievable rates for HD-FEC (asterisks) taken from Fig. 5 are also shown for comparison.

the GMI [6, eq. (4.55)], [9, eq. (13)], [10, eq. (25)]

$$\text{GMI} = \sum_{k=1}^m I(B_k; Y) = \sum_{k=1}^m I(B_k; L_k) \quad (13)$$

where the second equality holds if the LLRs are calculated via (3) and where $I(B_k; L_k)$ is the MI between the code bits and LLRs before FEC decoding. The three achievable rates considered above (MI, (12) and (13)) are schematically shown in Fig. 1.

Fig. 7 shows the GMI in (13) for different constellations. Similarly to the achievable rates for HD-FEC, the GMI curves cross each other (see black diamonds).⁶ Although this effect is less noticeable, the theoretical implications are the same: different SNRs call for different modulation sizes and FEC OH. In Fig. 7, we also show (with asterisks) the crossing points of the achievable rates for HD-FEC taken from Fig. 5. We do this to emphasize that if SD-FEC is considered instead of HD-FEC, different (higher) SNR thresholds are obtained. This is also visible in Fig. 8, where the optimum FEC OHs for SD-FEC are shown. The results in Figs. 7 and 8 have been recently experimentally studied in [46], [47].

C. Routing and Wavelength Assignment Problem

The routing and wavelength assignment problem is solved numerically as an integer linear programming (ILP) problem as described in [37, Section 4.1]. In particular, we maximize the network throughput in (10) (i.e., under a uniform traffic demand), where $C_{s,d}$ is assumed to be given by the capacity function C in (11). The ILP solution provides the number of active light paths and their routes between each node pair. From this, the total number of active light paths in the network is obtained. This solution also provides the SNR of each active light path.

⁶We again emphasize that this is only due to the fact a suboptimal receiver is considered. MI curves, on the other hand, do not cross each other for MQAM constellations, which has been known for many years (see, e.g., [44, Fig. 2], [45, Fig. 1]).

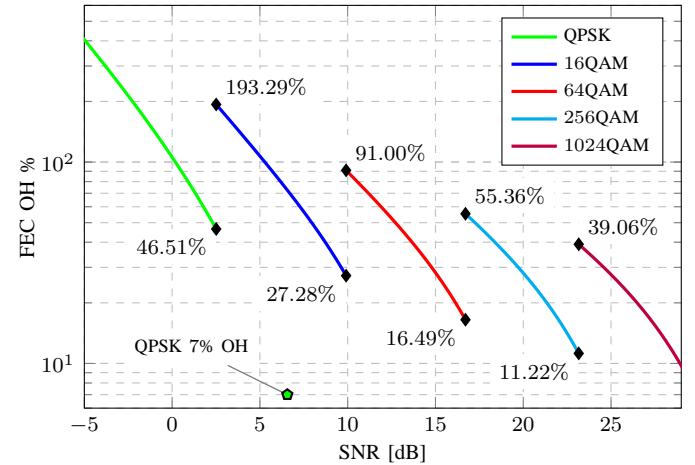


Fig. 8. OHs obtained from the achievable rates in Fig. 7 for a BW receiver with SD-FEC. The black diamonds show the SNR values where the modulation size should be changed and the corresponding FEC OH. The channel capacity in (11) is shown for comparison. The SNR required for QPSK with 7% FEC OH (SNR_{7%}) is also shown with a green pentagon.

The SNR values obtained by solving the ILP problem for the three network topologies in Figs. 2, 3, and 4 are shown in Fig. 9. The vertical bars show the number of transceivers that need to be installed to maximize the network throughput. The number of (two-way) transceivers for the DTG, NSF, and DTG networks are 1230, 1094, and 570, resp.

The vertical bars in Fig. 9 can be interpreted as the *distribution* of SNR across the network. By comparing these distributions, it is clear that the average SNRs across the network decreases as the size of the network increases. This is due to the presence of long links in continental and global networks (NSF and GB4). The SNR distributions in Fig. 9 also show that the *spread* of the SNR values is much larger for large networks. While for the DTG network, the variation in SNR is about 10 dB, for the GB4 network, this variation is about 20 dB.

Once the SNR values for the active light paths are found, the maximum network throughput can be calculated via (10). In particular, the SNRs of the routes shown in Fig. 9 are first grouped for each source destination pair. Then, the SNRs are “mapped” to throughputs via (11), and the value of $\sum_{r=1}^{R_{s,d}} C_{s,d}^{(r)}$ in (10) is obtained. The resulting throughputs are 524, 278, and 88 Tbps, for the DTG, NSF, and GB4 networks, resp. These throughput values are shown in the first columns of Table II together with the number of transceivers for each network.

D. Ideal FEC

The results in the previous section assume all the transceivers can achieve the capacity of the AWGN channel. This is never the case in practice as it requires the use of continuous constellations. In this section we consider the case where all transceivers can choose any of the MQAM constellations considered in this paper. Although more practically relevant, we assume the code rate R_c can be adjusted continuously, which is again never the case in practice. Nevertheless, the results in this section can be used

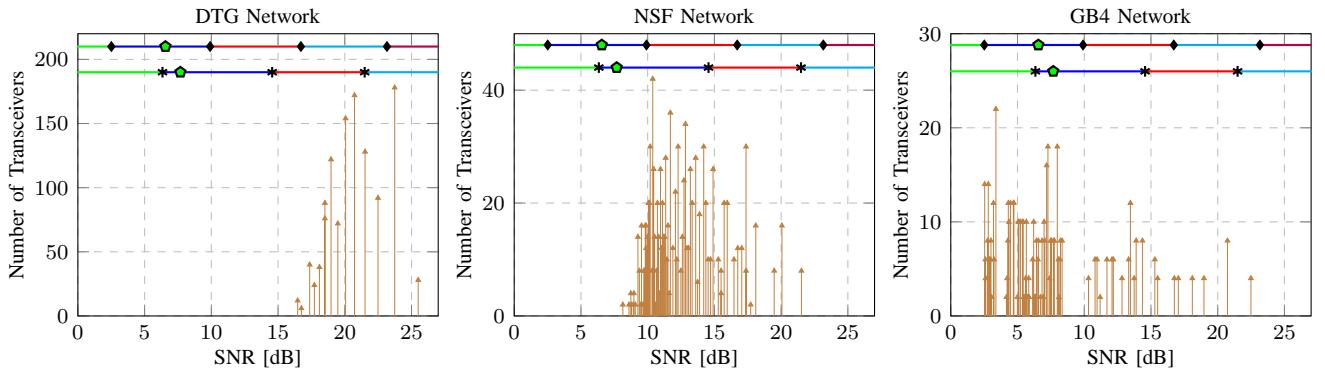


Fig. 9. Number of transceivers (vertical lines with triangles at the top) for all SNR values for the three network topologies in Figs. 2, 3, and 4. These values are obtained by assuming all transceivers can achieve the capacity C in (11). The horizontal lines show the SNR ranges in which different formats should be used. As in Figs. 6 and 8, asterisks and diamonds represent the SNR thresholds for HD- and SD-FEC, resp. The SNR thresholds for QPSK with 7% OH are also shown with green pentagons.

TABLE II
OPTIMUM NUMBER OF TRANSCEIVERS AND NETWORK THROUGHPUT Θ (IN Tbps) UNDER IDEALISTIC ASSUMPTIONS ON CODING AND MODULATION.

Network	Nodes N	Transceivers	Network Throughput Θ (in [Tbps])		
			Maximum (Capacity-based)	Ideal HD-FEC	Ideal SD-FEC
DTG	9	1230	524	431	488
NSFNET	14	1094	278	217	255
GB4	12	570	88	64	81

to estimate the penalty caused by the use of discrete (and square) constellations.

The selection of code rate and modulation format is assumed to be based on the achievable rates discussed in Sec. III-B. This idea is shown schematically in Fig. 9, where horizontal lines with different colors are included. These lines show the SNR ranges where different modulation formats should be used (lines with diamonds for SD-FEC and lines with stars for HD-FEC) and are obtained from Figs. 8 and 6.

To obtain the throughput achieved by ideal HD- and SD-FEC, we again use (10) and follow similar steps to those used in Sec. III-C. Namely, the SNRs of the routes (shown in Fig. 9) are first grouped for each source destination pair and the SNRs are then “mapped” to throughputs via the achievable rates discussed in Sec. III-B. The network throughputs obtained for HD-FEC, are 431, 217, and 64 Tbps, for the DTG, NSF, and GB4 networks, resp. For SD-FEC, these values become 488, 255, and 81 Tbps. These throughput results are shown in the last two columns of Table II.

The results in Table II show that, when compared to the maximum throughput obtained via the AWGN capacity assumption (fourth column in Table II), the use of ideal SD-FEC causes a relative throughput decrease of 9%, 8%, and 7%, for the DTG, NSF, and GB4 networks, resp. This indicates a relatively constant loss across different network topologies. On the other hand, the use of ideal HD-FEC causes relative losses of 18%, 22%, and 27%. These results show an increasing loss as the network size increases, which in turn shows the importance on considering SD-FEC for large networks. We conjecture that these increasing losses are due to the different shape of the “envelopes” of the crossing achievable rates in Figs. 5 and 7.

When compared to SD-FEC, HD-FEC codes are typically low complexity and low latency. On the other hand, for the same SNR, HD-FEC codes need higher OH to operate error free, which causes a throughput loss. The relative throughput losses are approximately 12%, 15%, and 20% for the DTG, NSF, and GB4 networks, resp. This indicates that low complexity and latency can be traded by a 10 – 20% loss in throughput and that the use SD-FEC becomes more and more important as the network size increases.

IV. PRACTICAL SCHEMES

Due to the continuous code rate assumption, the throughputs in the last two columns of Table II are only upper bounds that cannot be achieved in practice. In this section we discuss practically relevant alternatives.

A. QPSK with 7% FEC OH

Probably the simplest (and most popular) alternative in terms of coding and modulation for an optical network is to consider QPSK and a fixed FEC OH of 7% across the network. In this case, if the SNR of a given route is below the required SNR for QPSK with 7% OH, the route will not be used. If the SNR is above the threshold, then the available throughput in the r th route $C_{s,d}^{(r)}$ will be given by the SE in (1) times the symbol rate. The total available throughput (in [Tbps]) in (7) is then given by

$$C_{s,d}^{(r)} = \begin{cases} \frac{0.128}{1.07}, & \text{if } \text{SNR}_{s,d}^{(r)} \geq \text{SNR}_{7\%} \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where $\text{SNR}_{s,d}^{(r)}$ is the SNR of the r th route and $\text{SNR}_{7\%}$ is the SNR required for QPSK with 7% OH (shown with green pentagons in Figs. 6, 8, and 9).

The total network throughput in (10) (in Tbps) is then

$$\Theta = N(N-1) \cdot \frac{0.128}{1.07} \min_{\substack{s,d \in \mathcal{N} \\ d \neq s}} \sum_{r=1}^{R_{s,d}} I_{[\text{SNR}_{s,d}^{(r)} \geq \text{SNR}_{7\%}]} \quad (15)$$

where $I_{[\nu]}$ is an indicator function, i.e., $I_{[\nu]} = 1$ if ν is true, and $I_{[\nu]} = 0$ otherwise.

The SNR threshold $\text{SNR}_{7\%}$ in (15) is different for HD- and SD-FEC, and thus, the resulting throughputs might also be different. However, for both the DTG and NSF networks, all the route SNRs are above both thresholds, and thus, the total network throughput in (15) is

$$\Theta = N(N-1) \cdot \frac{0.128}{1.07} \min_{\substack{s,d \in \mathcal{N} \\ d \neq s}} \{R_{s,d}\}. \quad (16)$$

The minimum number of routes for any source destination pair for the NTG and NSF networks are 14 and 4, resp., which combined with the number of nodes in the network, give throughputs of 120 and 87 Tbps.

For QPSK and 7% FEC OH, SD offers a theoretical maximum sensitivity increase of about 1.15 dB⁷. However, for the DTG and NSF networks, there is no difference between HD and SD in terms of network throughput as the minimum SNR of all routes is above the SNR threshold. This result highlights the fact that under these conditions and traffic assumptions, upgrading all transceivers from HD to SD FEC without changing the OH might not bring any benefit for small networks.

When the GB4 network is considered, however, most of the routes are in fact below the SNR required for QPSK with 7% OH (see Fig. 9). For this case, the network throughput given by (15) is in fact zero. This can be intuitively explained by the fact that there are node pairs that are very far apart, and thus, the uniform throughput constraint and full network connectivity cannot be satisfied.

B. One M Schemes

As an alternative to the QPSK with 7% FEC OH approach, in this section we consider two approaches, both of them based on the philosophy that only one modulation format should be implemented across the network.

The first scheme assumes all transceivers implement one modulation format and one code rate. We call this scheme $1R_c1M$. In this scheme, the code rate is chosen such that the network throughput (based on achievable rates) is maximized. The results obtained for $1R_c1M$ are shown in Fig. 10 with red triangles for HD-FEC and with blue triangles for SD-FEC. The top figure shows network throughput, while the bottom ones show optimum code rates. The different networks under consideration are shown from left to right. In the throughput results, we also include the ideal network throughputs in Table II (solid horizontal lines) as well as the results obtained

⁷This can be obtained by comparing the green pentagons for HD and SD FEC in Figs. 6 and 8.

with QPSK and 7% FEC OH from Sec. IV-A (green pentagons).

The results in Fig. 10 show that for HD-FEC and $1R_c1M$, there is always a modulation format that is optimum: $M = 64$ for DTG, $M = 16$ for NSF, and $M = 4$ for GB4. Nevertheless, for both HD- and SD-FEC, the gains obtained (with respect to QPSK with 7% OH) by using $1R_c1M$ are quite large. Interestingly, the optimum code rates for HD-FEC and $1R_c1M$ are $R_c \approx 0.8$ for DTG, $R_c \approx 0.62$ for NSF, and $R_c \approx 0.56$ for GB4 (24%, 61% and 79% FEC OH, resp.). This indicates a clear potential benefit of using large FEC OH when HD-FEC is used in a network context and where the modulation format is fixed across the network.

When it comes to $1R_c1M$ with SD-FEC, the optimality of a given modulation format is less clear, as the throughput curves in this case do not have a clear peak. This is due to the fact that the crossings between the GMI curves is not as pronounced as the crossings for HD-FEC. Nevertheless, by observing the trend of the curves, a good compromise would be to choose the same modulation formats as for HD-FEC, i.e., $M = 64$ for DTG, $M = 16$ for NSF, and $M = 4$ for GB4. In this case, the corresponding optimum code rates are $R_c \approx 0.89$, $R_c \approx 0.75$, and $R_c \approx 0.68$ (12%, 33% and 47% FEC OH, resp.). In general, the optimum code rates in this case are slightly higher than the ones for HD-FEC. The gains of SD-FEC over HD-FEC in this case are approximately 50 Tbps for the DTG and NSF networks. On the other hand, only small gains are observed for the GB4 network.

As mentioned before, the throughput for QPSK with 7% FEC OH for the GB4 network is zero. On the other hand, for this network, $1R_c1M$ gives throughputs of about 20 Tbps. This is obtained by using QPSK and an increased OH. This result highlights the need for considering high FEC OH in large networks.

The second scheme we consider in this section is called $2R_c1M$. In this case, the transceivers are equipped with two code rates but only one modulation format. Again, the code rates are optimized so that the network throughput is maximized.

The results obtained for $2R_c1M$ are shown in Fig. 10 with circles (red for HD-FEC and blue for SD-FEC). These results show that, regardless of the network under consideration, half of the gap between the ideal FEC limit (horizontal lines) and the throughput obtained by $1R_c1M$ can be harvested by using an extra code rate. These results highlight the advantage of adapting the code rate to the variable channel conditions across the network. These results also suggest that a good complexity-performance tradeoff is obtained by using one modulation format and two code rates.

C. Variable M Schemes

In this section we consider a design alternative where the modulation format is variable. In particular, we assume all the transceivers can choose any modulation format in between QPSK and a given maximum value of M , which we denote by \hat{M} . For example, if $\hat{M} = 64$, then the transceivers can choose from $M = 4$, $M = 16$ and $M = 64$.

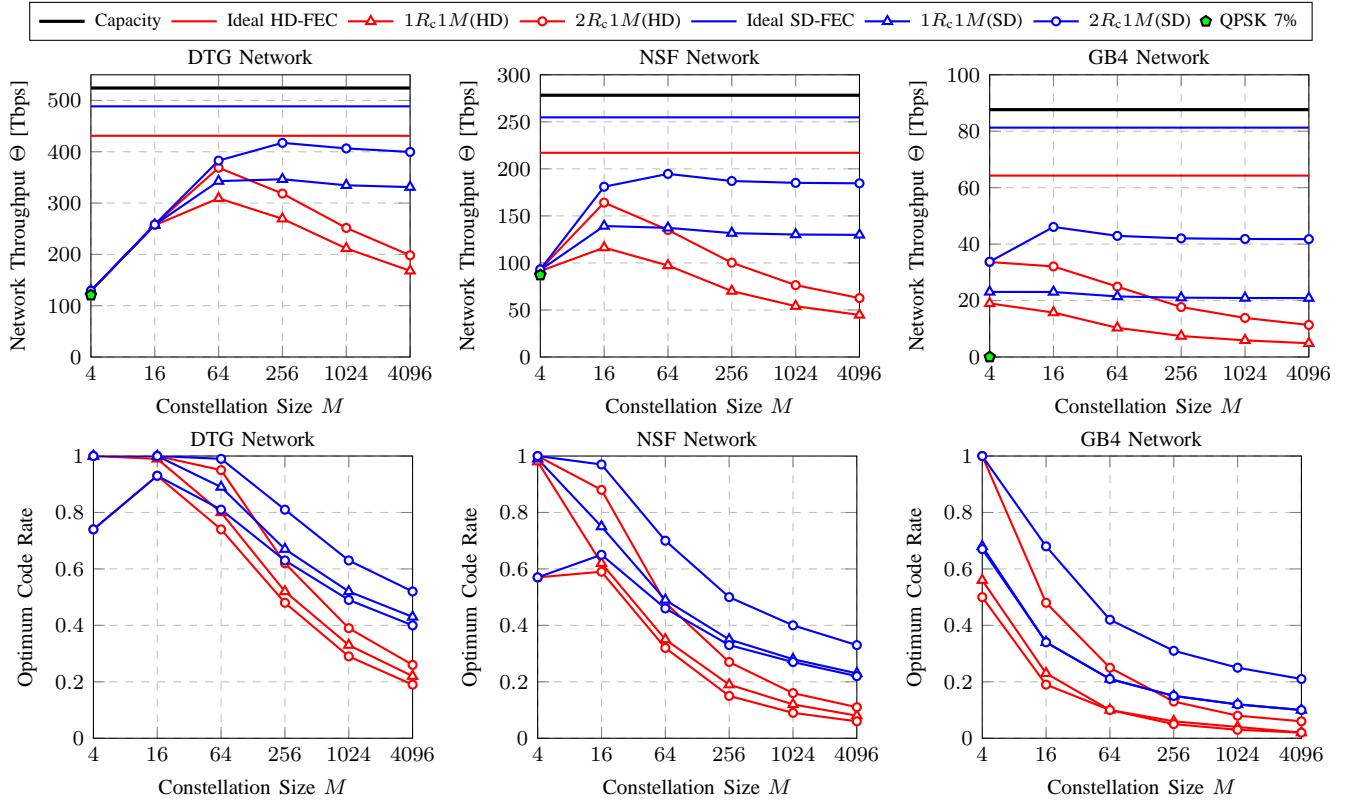


Fig. 10. Throughput (top row) and optimum code rate (bottom row) for one M schemes, for both HD-FEC (red) and SD-FEC (blue), and for the three network topologies in Figs. 2, 3, and 4. The throughput results obtained with QPSK and 7% FEC OH in Sec. IV-A are also shown with green pentagons.

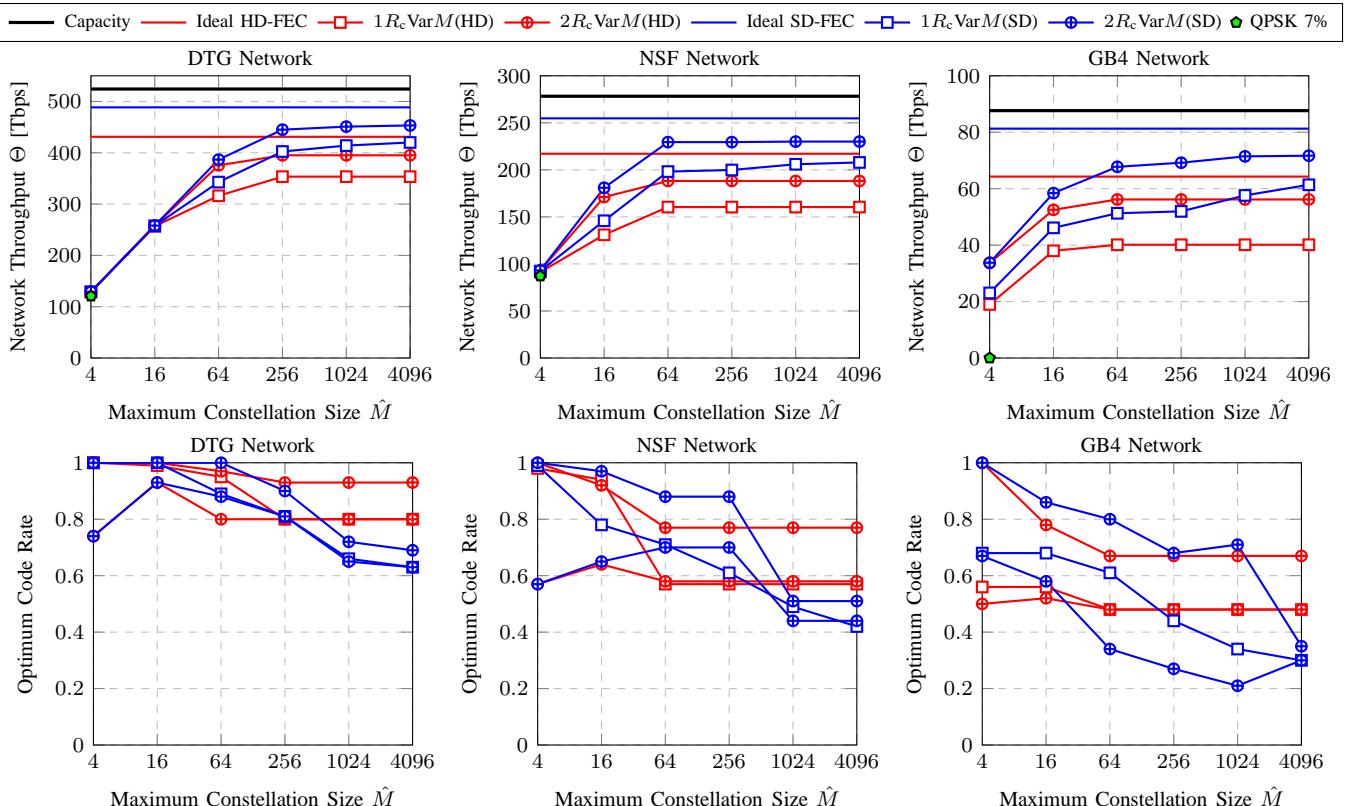


Fig. 11. Throughput (top row) and optimum code rate (bottom row) for variable M schemes, for both HD-FEC (red) and SD-FEC (blue), and for the three network topologies in Figs. 2, 3, and 4. The throughput results obtained with QPSK and 7% FEC OH in Sec. IV-A are also shown with green pentagons.



Fig. 12. Percentage of transceivers using different modulation formats for the optimal HD-FEC solution (for $1R_cVarM$) and the three network topologies. The results are shown for different maximum constellation sizes \hat{M} .

The first scheme we consider assumes only one code rate is implemented across the network, and that the code rate is optimized to maximize the network throughput. We denote this scheme $1R_cVarM$. The results obtained using this scheme are shown in Fig. 11 (squares) and indicate that, in general, constellation sizes beyond 256QAM give little throughput increases (the throughput curves flatten out for large values of \hat{M}).

Figs. 12 and 13 show the percentage of transceivers using different modulation formats for $1R_cVarM$, for HD- and SD-FEC, resp., and for $\hat{M} = 16, 64, 256$. The value in the middle of each chart is the total throughput obtained. The general trend in these results is that QPSK (green) is only useful in the very large network (GB4). These results also show that, considering 256QAM (light blue) gives very little throughput increases for the NSF and GB4 networks, for both HD- and SD-FEC. On the other hand, 256QAM gives a relevant throughput increase for the DTG network.

To compare the throughput contribution of different modulation formats for $1R_cVarM$, we show in Fig. 14 the obtained throughputs for both HD- and SD-FEC. Apart from showing the relative contributions, this figure also shows how the contributions change when SD-FEC is considered instead of HD-FEC. The throughput gains due to SD-FEC are also clearly visible.

The results in Figs. 11, 12, 13, and 14 show that $1R_cVarM$ gives similar throughput results to those obtained by $2R_c1M$ (two code rates and one modulation format) shown in Fig. 10. This indicates that similar (large) gains, can be

obtained by having either multiple code rates or multiple modulation formats.

The second scheme we consider in this section is one where two code rates are implemented across the network (and the modulation can be varied too). We call this scheme $2R_cVarM$. The results are shown with plus-circles in Fig. 11. These results indicate that including a second code rate gives a clear advantage with respect to $1R_cVarM$ (squares). This gain is particularly visible for large values of \hat{M} and large networks. In particular, for the GB4 network and $\hat{M} = 256$, the gains are approximately 15 Tbps for both HD- and SD-FEC.

We conclude this section by comparing the optimal code rates of two-rate versus the one-rate schemes. In particular, we observe from Figs. 10 and 11 that the smallest code rate for the two-rate schemes is always quite close to the optimal code rate of the corresponding one-rate scheme. The intuition behind this is that when the transmitters are equipped with two code rates, one low code rate can be used for the worst performing connection, while the other rate is used to increase the overall network throughput.

V. CONCLUSIONS

Optimal constellation sizes and FEC overheads for optical networks were studied. Joint optimization of the constellation and FEC OHs was shown to yield large gains in terms of overall network throughput. The optimal values were shown to be dependent on the SNR distributions within a network. Two code rates and a single constellation (which varies as a function of the network size) gave a good throughput-complexity tradeoff.

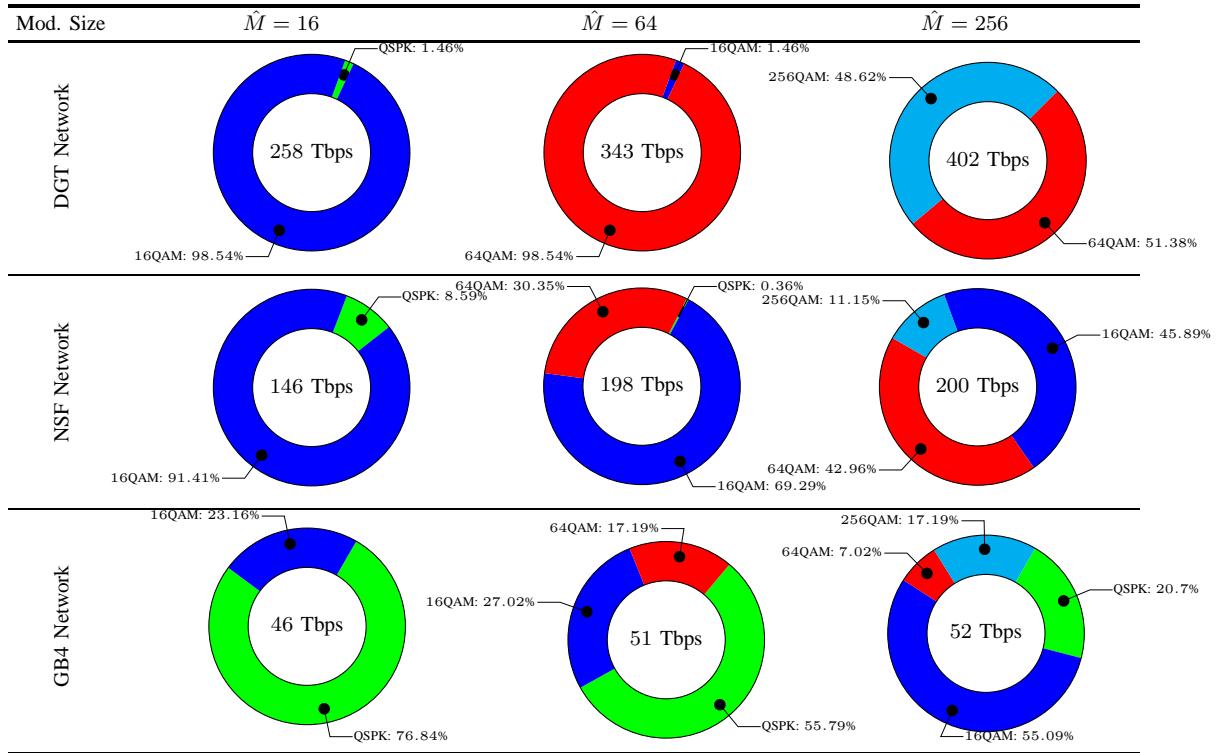


Fig. 13. Percentage of transceivers using different modulation formats for the optimal SD-FEC solution (for $1R_c \text{Var}M$) and the three network topologies. The results are shown for different maximum constellation sizes \hat{M} .

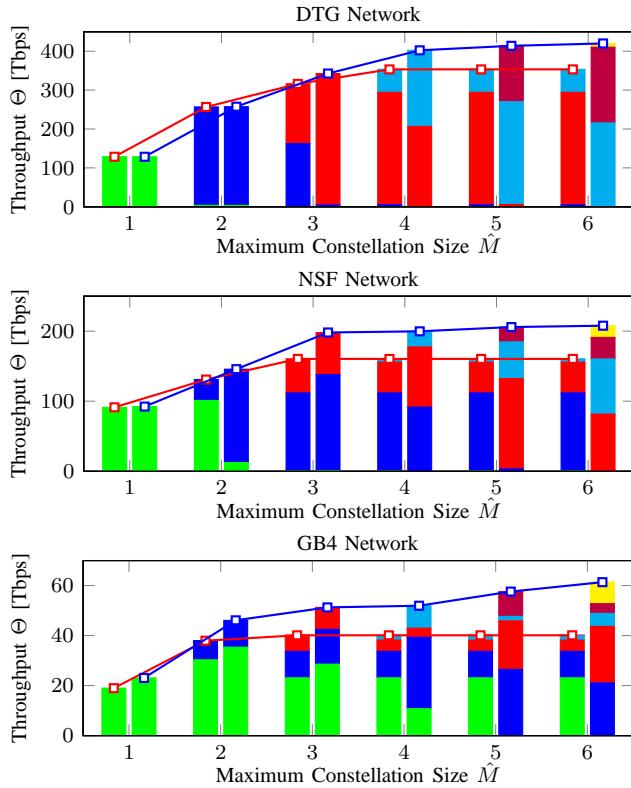


Fig. 14. Network throughput for $1R_c \text{Var}M$ for HD-FEC (red lines with squares) and SD-FEC (blue lines with squares) as a function of the maximum constellation size for the three network topologies in Figs. 2, 3, and 4. The colors represent the throughput contribution from the different modulation sizes. The lines with markers are the network throughput shown in Fig. 11.

In this paper we studied the problem from the point of view of largest achievable rates. Practical FEC implementations, however, will operate a few decibels (or fractions of decibels) away from these achievable rates. Nevertheless, if these penalties are known a priori, the methodology used in this paper can be straightforwardly used to consider practical codes. This is left for future investigation.

REFERENCES

- [1] A. Chralyvy, "Plenary paper: The coming capacity crunch," in *Proc. European Conference on Optical Communication (ECOC)*, Vienna, Austria, Sep. 2009.
- [2] D. J. Richardson, "Filling the light pipe," *Science*, vol. 330, no. 6002, pp. 327–1328, Oct. 2010.
- [3] M. Jinno, B. Koziicki, H. Takara, A. Watanabe, Y. Sone, T. Tanaka, and A. Hirano, "Distance-adaptive spectrum resource allocation in spectrum-sliced elastic optical path network," *IEEE Commun. Mag.*, vol. 48, no. 8, pp. 138–145, Aug. 2010.
- [4] Y. Lia, H. Daia, G. Shena, and S. K. Bose, "Adaptive FEC-based lightpath routing and wavelength assignment in WDM optical networks," *Optical Switching and Networking*, vol. 14, no. 3, pp. 241–249, Aug. 2014.
- [5] S. Gringeri, N. Bitar, and T. J. Xia, "Extending software defined network principles to include optical transport," *IEEE Commun. Mag.*, vol. 52, no. 3, pp. 32–40, Mar. 2013.
- [6] L. Szczecinski and A. Alvarado, *Bit-Interleaved Coded Modulation: Fundamentals, Analysis and Design*. John Wiley & Sons, 2015.
- [7] H. Bülow, Ü. Abay, A. Schenk, and J. B. Huber, "Coded modulation of polarization- and space-multiplexed signals," in *Asia Communications and Photonics Conference and Exhibition (ACP)*, Shanghai, China, Nov. 2011.
- [8] D. S. Millar, T. Koike-Akino, R. Maher, D. Lavery, M. Paskov, K. Kojima, K. Parsons, B. C. Thomsen, S. J. Savory, and P. Bayvel, "Experimental demonstration of 24-dimensional extended Golay coded modulation with LDPC," in *Proc. Optical Fiber Communication Conference (OFC)*, San Francisco, CA, Mar. 2014.

- [9] A. Alvarado and E. Agrell, "Four-dimensional coded modulation with bit-wise decoders for future optical communications," *J. Lightw. Technol.*, vol. 33, no. 10, pp. 1993–2003, May 2015.
- [10] A. Alvarado, E. A. D. Lavery, R. Maher, and P. Bayvel, "Replacing the soft FEC limit paradigm in the design of optical communication systems," *J. Lightw. Technol.*, 2015 (to appear).
- [11] A. Nag, M. Tornatore, and M. Mukherjee, "Optical network design with mixed line rates and multiple modulation formats," *J. Lightw. Technol.*, vol. 28, no. 4, pp. 466–475, Feb. 2010.
- [12] G.-H. Gho, L. Klak, and J. M. Kahn, "Rate-adaptive coding for optical fiber transmission systems," *J. Lightw. Technol.*, vol. 29, no. 2, pp. 222–233, Jan. 2011.
- [13] G.-H. Gho and J. M. Kahn, "Rate-adaptive modulation and low-density parity-check coding for optical fiber transmission systems," *J. Opt. Commun. Netw.*, vol. 4, no. 10, pp. 760–768, Oct. 2012.
- [14] D. A. A. Mello, A. N. Barreto, T. C. de Lima, T. F. Portela, L. Beygi, and J. M. Kahn, "Optical networking with variable-code-rate transceivers," *J. Lightw. Technol.*, vol. 32, no. 2, pp. 257–266, Jan. 2014.
- [15] S. J. Savory, "Congestion aware routing in nonlinear elastic optical networks," *IEEE Photon. Technol. Lett.*, vol. 26, no. 10, pp. 1057–1060, May 2014.
- [16] N. Sambo, G. Meloni, F. Cugini, A. D'Errico, L. Potì, P. Iovanna, and P. Castoldi, "Routing, code, and spectrum assignment (RCSA) in elastic optical networks," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015.
- [17] N. Sambo, G. Meloni, F. Paolucci, F. Cugini, M. Secondini, F. Fresi, L. Potì, and P. Castoldi, "Programmable transponder, code and differentiated filter configuration in elastic optical networks," *J. Lightw. Technol.*, vol. 32, no. 11, pp. 2079–2086, June 2014.
- [18] K. Christodoulopoulos, I. Tomkos, and E. Varvarigos, "Elastic bandwidth allocation in flexible OFDM-Based optical networks," *J. Lightw. Technol.*, vol. 29, no. 9, pp. 1354–1366, May 2011.
- [19] R. Schmogrow, D. Hillerkuss, M. Dreschmann, M. Huebner, M. Winter, J. Meyer, B. Nebendahl, C. Koos, J. Becker, W. Freude, and J. Leuthold, "Real-time software-defined multiformat transmitter generating 64QAM at 28 GBd," *IEEE Photon. Technol. Lett.*, vol. 22, no. 21, pp. 1601–1603, Nov. 2010.
- [20] K. Roberts and C. Laperle, "Flexible transceivers," in *Proc. European Conference on Optical Communication (ECOC)*, Amsterdam, The Netherlands, Sep. 2012.
- [21] H. Y. Choi, T. Tsuritani, and I. Morita, "Ber-adaptive flexible-format transmitter for elastic optical networks," *Opt. Express*, vol. 20, no. 17, pp. 18 652–18 658, Aug. 2012.
- [22] B. T. Teipen, M. H. E. an K. Grobe, and J.-P. Elbers, "Adaptive data rates for flexible transceivers in optical networks," *Journal of Networks*, vol. 17, no. 5, pp. 776–782, May 2012.
- [23] H. Y. Choi, L. L. and Takehiro Tsuritani, and I. Morita, "Demonstration of BER-Adaptive WSON employing flexible transmitter/receiver with an extended openflow-based control plane," *IEEE Photon. Technol. Lett.*, vol. 25, no. 2, pp. 119–121, Jan. 2013.
- [24] N. Sambo, F. Paolucci, F. Cugini, M. Secondini, L. Potì, G. Berrettini, G. Meloni, F. Fresi, G. Bottari, and P. Castoldi, "Software defined code-rate-adaptive terabit/s based on time-frequency packing," in *Proc. Optical Fiber Communication Conference (OFC)*, Anaheim, CA, Mar. 2013.
- [25] A. Splett, C. Kurtzke, and K. Petermann, "Ultimate transmission capacity of amplified optical fiber communication systems taking into account fiber nonlinearities," in *Proc. European Conference on Optical Communication (ECOC)*, Montreux, Switzerland, Sept. 1993.
- [26] E. Grellier and A. A. Bononi, "Quality parameter for coherent transmissions with Gaussian-distributed nonlinear noise," *Opt. Express*, vol. 19, no. 13, pp. 12 781–12 788, June 2011.
- [27] P. Poggioletti, A. Carena, V. Curri, G. Bosco, and F. Forghieri, "Analytical modeling of nonlinear propagation in uncompensated optical transmission links," *IEEE Photon. Technol. Lett.*, vol. 23, no. 11, pp. 742–744, June 2011.
- [28] L. Beygi, E. Agrell, P. Johannsson, M. Karlsson, and H. Wymeersch, "A discrete-time model for uncompensated single-channel fiber-optical links," *IEEE Trans. Commun.*, vol. 60, no. 11, pp. 3440–3450, Nov. 2012.
- [29] P. Johannsson and M. Karlsson, "Perturbation analysis of nonlinear propagation in a strongly dispersive optical communication system," *J. Lightw. Technol.*, vol. 31, no. 8, pp. 1273–1282, Apr. 2013.
- [30] P. Poggioletti, G. Bosco, A. Carena, V. Curri, Y. Jiand, and F. Forghieri, "The GN-model of fiber non-linear propagation and its applications," *J. Lightw. Technol.*, vol. 32, no. 4, pp. 694–721, Oct. 2014, (Invited Tutorial).
- [31] P. Johannsson and E. Agrell, "Modeling of nonlinear signal distortion in fiber-optic networks," *J. Lightw. Technol.*, vol. 32, no. 23, pp. 4544–4552, Dec. 2014.
- [32] J. Zhao, H. Wymeersch, and E. Agrell, "Nonlinear impairment aware resource allocation in elastic optical networks," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015.
- [33] L. Yan, E. Agrell, and H. Wymeersch, "Resource allocation in nonlinear flexible-grid fiber-optic networks," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015.
- [34] —, "Link-level resource allocation for flexible-grid nonlinear fiber-optic communication systems," *IEEE Photon. Technol. Lett.*, vol. 27, no. 12, pp. 1250–1253, June 2015.
- [35] D. J. Ives, P. Bayvel, and S. J. Savory, "Adapting transmitter power and modulation format to improve optical network performance utilizing the Gaussian noise model of nonlinear impairments," *J. Lightw. Technol.*, vol. 32, no. 21, pp. 4087–4096, Nov. 2014.
- [36] —, "Physical layer transmitter and routing optimization to maximize the traffic throughput of a nonlinear optical mesh network," in *International Conference on Optical Network Design and Modeling*, Stockholm, Sweden, May 2014.
- [37] —, "Routing, modulation, spectrum and launch power assignment to maximize the traffic throughput of a nonlinear optical mesh network," *Photonic Network Communications*, vol. 29, no. 3, pp. 244–256, Mar. 2015.
- [38] —, "Assessment of options for utilizing SNR margin to increase network data throughput," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015.
- [39] A. Alvarado, D. J. Ives, S. J. Savory, and P. Bayvel, "On optimal modulation and fec overhead for future optical networks," in *Proc. Optical Fiber Communication Conference (OFC)*, Los Angeles, CA, Mar. 2015.
- [40] S. Makovejs, C. Behrens, R.-P. Braun, S. Ten, C. Towery, I. Roudas, K. Koreshkov, T. Nath, and A. Gladisch, "Impact of adaptive-rate transponders and fiber attributes on the achievable capacity," *Journal of Optical Communications and Networking*, vol. 7, no. 3, pp. 172–175, Mar. 2015.
- [41] S. Jain, A. Kumar, S. Mandal, J. Ong, L. Poutievski, A. Singh, S. Venkata, J. Wanderer, J. Zhou, M. Zhu, J. Zolla, U. Hözlé, S. Stuart, and A. Vahdat, "B4: Experience with a globally-deployed software defined wan," *SIGCOMM Comput. Commun. Rev.*, vol. 43, no. 4, pp. 3–14, Aug. 2013.
- [42] T. Fehnberger, A. Alvarado, P. Bayvel, and N. Hanik, "On achievable rates for long-haul fiber-optic communications," *Opt. Express*, vol. 23, no. 7, pp. 9183–9191, Apr. 2015.
- [43] M. Ivanov, F. Brännström, A. Alvarado, and E. Agrell, "General BER expression for one-dimensional constellations," in *IEEE Global Telecommunications Conference (GLOBECOM)*, Anaheim, CA, Dec. 2012.
- [44] G. Ungerboeck, "Channel coding with multilevel/phase signals," *IEEE Trans. Inf. Theory*, vol. 28, no. 1, pp. 55–67, Jan. 1982.
- [45] G. D. Forney, Jr. and G. Ungerboeck, "Modulation and coding for linear Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2384–2415, Oct. 1998, (Invited Paper).
- [46] R. Maher, A. Alvarado, D. Lavery, and P. Bayvel, "Multi-channel DBP for reach enhancement of high capacity M-QAM super-channels," in *Signal Processing in Photonic Communications (SPPCOM)*, Boston, MA, June-July 2015.
- [47] —, "Modulation order and code rate optimisation for digital coherent transceivers using generalised mutual information," in *Proc. European Conference on Optical Communication (ECOC)*, Valencia, Spain, Sep. 2015.